

# ON THE LIMITING SPECTRAL DENSITY OF RANDOM MATRICES FILLED WITH STOCHASTIC PROCESSES

MATTHIAS LÖWE AND KRISTINA SCHUBERT

**ABSTRACT.** We discuss the limiting spectral density of real symmetric random matrices. Other than in standard random matrix theory the upper diagonal entries are not assumed to be independent, but we will fill them with the entries of a stochastic process. Under assumptions on this process, which are satisfied, e.g., by stationary Markov chains on finite sets, by stationary Gibbs measures on finite state spaces, or by Gaussian Markov processes, we show that the limiting spectral distribution depends on the way the matrix is filled with the stochastic process. If the filling is in a certain way compatible with the symmetry condition on the matrix, the limiting law of the empirical eigenvalue distribution is the well known semi-circle law. For other fillings we show that the semi-circle law cannot be the limiting spectral density.

## 1. INTRODUCTION

A central aspect in the study of random matrices with growing dimension is the analysis of their eigenvalue distribution. To introduce them let, for any  $N \in \mathbb{N}$ ,  $\{X_N(i, j), 1 \leq i \leq j \leq N\}$  be a real valued random field. Define the symmetric random  $N \times N$  matrix  $A_N$  by

$$A_N(i, j) = \frac{1}{\sqrt{N}} X_N(i, j), \quad 1 \leq i \leq j \leq N.$$

We will denote the (real) eigenvalues of  $A_N$  by  $\lambda_1^{(N)} \leq \lambda_2^{(N)} \leq \dots \leq \lambda_N^{(N)}$ . Let  $\mu_N$  be the empirical eigenvalue distribution, i.e.

$$\mu_N = \frac{1}{N} \sum_{k=1}^N \delta_{\lambda_k^{(N)}}.$$

In [20] and [21] Wigner proved that, if  $X_N(p, q), 1 \leq p \leq q \leq N$ , are independent random variables with expectation 0, and either Bernoulli or normally distributed with variance 1 for off-diagonal elements and variance 2 on the diagonal, the empirical eigenvalue distribution  $\mu_N$  converges weakly in probability to the so called semicircle distribution (or law), i.e. the probability distribution  $\nu$  on  $\mathbb{R}$  with density

$$\nu(dx) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{|x| \leq 2}. \quad (1.1)$$

---

2010 *Mathematics Subject Classification.* 60B20; 60F05.

*Key words and phrases.* random matrix, dependent entries, stochastic processes, Wigner's semi-circle law.

Arnold [2] generalized this result by showing that the convergence to the semicircle law is also true, if one replaces the Bernoulli or Gaussian distribution by the assumption that the random variables are independent and identically distributed (i.i.d.) and have a finite fourth moment. This observation can be considered the starting point of a collection of results showing the universality of the semicircle law. For example, also the identical distribution of the random variables may be replaced by a weaker assumptions (see e.g. [5]). Moreover, Erdős et al. [6]) and Tao and Vu [19] observed that the convergence of the empirical eigenvalue distribution towards the semi-circle law holds under general assumptions in a local sense. The interested reader is referred to [18] for more results on universality of the semi-circle law, in particular the so called “four moment theorem”.

Another generalizations of Wigners semi-circle law deals with matrix ensembles with entries realized according to weighted Haar measures on classical (e.g. orthogonal, unitary, symplectic) groups. Such results lay a bridge between classical and non-commutative probability (see e.g. [12], or the recommendable book by Anderson, Guionnet, and Zeitouni [1]).

A different approach to universality was taken in [16], [11] and [9]. In all these articles random matrices with correlated entries are studied. While the first of these papers imposes the condition that the number of correlated matrix entries satisfies a certain bound, the second assumes that the entries of the random matrix satisfy a certain martingale condition. Moreover, [11] establishes a Stein approach to the semicircle law. In [9] it is shown that, if the diagonals of  $X_N$  are independent and the correlation between elements along a diagonal decays sufficiently quickly, again the limiting spectral distribution is the semi-circle law. This result has to be compared with the situation in [8] and [7], where the diagonals are still independent, but the random variables along a diagonal are exchangeable. Here again, for weak correlations one finds the semicircle as the limiting spectral density for the eigenvalues. However, for stronger correlations the limiting spectral measure is a free convolution of the semicircle law with the limit measure for the eigenvalues of a normalized Toeplitz matrix, which was found in [3]. This means that one observes a kind of phase transition in the limiting spectral density. A similar results for matrices of Hankel type was shown in [17]. In [13] the authors generalize the conditions from [8] by assuming that now the entire upper diagonal entries of the matrix are exchangeable. Again, if the correlations decay quickly enough, the limiting spectral distribution is the semicircle law. This is, to the best of our knowledge, one of the rare occasions that the semi-circle law has been shown for a random matrix without any independence assumptions.

The results in [9] and [13] immediately raise the following question: What is the limiting spectral distribution of a random matrix, if all the upper diagonal entries are correlated but not exchangeable. Considering the matrix entries as realizations of a stochastic process, a natural assumption is again that the correlations decay quickly, e.g. like the correlations in a Markov chain. This

is exactly the problem we are going to attack in the present paper. In a nutshell, the result is that under natural assumptions, e.g. centered entries if the stochastic process is a Markov chain, the semicircle law will be the limiting spectral distribution. However, this results depends on the way the matrix is filled with the realization of the stochastic process. We will need a condition that is e.g. satisfied, if the stochastic process follows the diagonals of the random matrix. We will also show that there are Markov chains (even on  $\{-1, +1\}$ ) that, if filled into a matrix row- or column-wise, lead to a spectral distribution that does not converge to the semicircle law.

We organize this paper in the following way: The second section contains our basic definitions together with the central results. In Section 3 we focus on the assumptions for our results. We show that suitably chosen Markov chains, Gibbs measures and Gaussian Markov processes are examples for the stochastic processes that appear in our main theorem and we consider the filling, where the stochastic process follows the diagonals of the matrix. Section 4 contains the proofs of our results.

## 2. RANDOM MATRICES WITH ENTRIES FROM A STOCHASTIC PROCESS

Given a stochastic process  $(Z_n)_n$ , there are, of course, various ways to fill them into a (symmetric) random matrix. The following definition formalizes these ways. It also introduces a distance on the indices of the matrix, which will enable us to formulate a condition, under which the spectral distribution of the corresponding matrices converges to the semicircle law. This distance is induced by the filling of the matrix.

### Definition 2.1.

- (a) A symmetric matrix  $X_N \in \mathbb{R}^{N \times N}$  is generated by a stochastic process  $(Z_n)_n$  and a mapping  $\varphi_N$ , if
  - (i)  $\varphi_N : \{1, \dots, N(N+1)/2\} \rightarrow \{(i, j) \in \{1, \dots, N\}^2 : i \leq j\}$  is bijective,
  - (ii) the matrix entries (in the upper triangular part of the matrix) are given by  $(Z_n)_n$  via

$$X_N(i, j) = Z_{\varphi_N^{(-1)}(i, j)}, \quad 1 \leq i \leq j \leq N.$$

- (b) We measure the distance between two matrix indices  $(i, j)$  and  $(i', j')$  of a  $N \times N$  matrix in terms of their distance in the stochastic process, i.e. we set

$$\|(i, j) - (i', j')\|_{\varphi_N} := |\varphi_N^{(-1)}(i, j) - \varphi_N^{(-1)}(i', j')|, \quad i \leq j, i' \leq j'.$$

If  $i > j$  or  $i' > j'$ , we use  $(j, i)$  resp.  $(j', i')$  instead to calculate the distance.

- (c) We call matrix indices  $(i, j), (i', j') \in \{1, \dots, N\}^2$ ,  $i \leq j$ ,  $i' \leq j'$  (horizontal or vertical, respectively) neighbors if
  - $i = i'$  and  $j \in \{j' + 1, j' - 1\}$  or

- $j = j'$  and  $i \in \{i' + 1, i' - 1\}$ , respectively.

Considering the path through the upper triangular part of a matrix prescribed by the mapping  $\varphi_N$ , we denote the number of steps to the horizontal or vertical neighboring matrix entry by

$$J(\varphi_N) := \# \left\{ 1 \leq i < \frac{N(N+1)}{2} : \varphi_N(i) \text{ and } \varphi_N(i+1) \text{ are neighbors} \right\}.$$

In our main theorem we study the convergence of the empirical eigenvalue distribution of matrices generated by a stochastic process  $(Z_n)_n$  with correlations that decay quickly and a mapping  $\varphi_N$ . We provide conditions on the mapping  $\varphi_N$  and the stochastic process  $(Z_n)_n$  under which  $\mu_N$  converges weakly in probability to Wigner's semicircle law  $\nu$ . These conditions are:

**Assumption 2.2.** Assume that the mapping  $\varphi_N$  satisfies

$$\forall i, j, n : \# \{x \in \{1, \dots, N\} : \|(i, x) - (x, j)\|_{\varphi_N} = n\} = o(N).$$

**Assumption 2.3.** Let  $(Z_n)_{n \in \mathbb{N}}$  be a stochastic process. Assume the following:

- (i) for all  $i \in \mathbb{N}$  we have  $\mathbb{E}(Z_i^2) = 1$ ,
- (ii) for all odd numbers  $k$  and all  $i_1, \dots, i_k$  we have that

$$\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k}) = 0, \quad (2.2)$$

- (iii) for all even  $k$  there are constants  $C > 0$  and  $\beta \in [0, 1)$  such that for all  $i_1 \leq i_2 \leq \dots \leq i_k$  with  $n_j := i_{j+1} - i_j$  we have

$$|\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k})| \leq C \beta^{n_1 + n_3 + \dots + n_{k-1}}, \quad (2.3)$$

for  $j_1 \leq \dots \leq j_k$  with  $d := \min_{n, m \in \{1, \dots, k\}} |i_n - j_m|$  we have

$$|\mathbb{E}(Z_{i_1} \dots Z_{i_k} Z_{j_1} \dots Z_{j_k}) - \mathbb{E}(Z_{i_1} \dots Z_{i_k}) \mathbb{E}(Z_{j_1} \dots Z_{j_k})| \leq C \beta^d, \quad (2.4)$$

and for  $d' := \min_{j=1,3,5,\dots} i_{j+2} - i_j$  we have

$$\left| \mathbb{E}(Z_{i_1}^2 Z_{i_3}^2 \dots Z_{i_{k-1}}^2) - 1 \right| \leq C \beta^{d'}. \quad (2.5)$$

*Remark 2.4.* If, instead of (2.4) we assume that for all  $k$  there exists  $C > 0$  and  $\beta \in [0, 1)$  such that for all  $i_1 \leq i_2 \leq \dots \leq i_k$  and  $j_1 \leq \dots \leq j_l$  with  $l \leq k$  and  $d := \min_{\substack{n \in \{1, \dots, k\} \\ m \in \{1, \dots, l\}}} |i_n - j_m|$  we have

$$|\mathbb{E}(Z_{i_1} \dots Z_{i_k} Z_{j_1} \dots Z_{j_l}) - \mathbb{E}(Z_{i_1} \dots Z_{i_k}) \mathbb{E}(Z_{j_1} \dots Z_{j_l})| \leq C \beta^d,$$

the estimate (2.5) follows automatically. However, all these assumptions are met by our examples.

With these assumptions our main theorem reads as follows.

**Theorem 2.5.** Let  $X_N \in \mathbb{R}^{N \times N}$  be symmetric matrices generated by a stochastic process  $(Z_n)_n$  that fulfills Assumption 2.3 and a filling  $\varphi_N$  that satisfies Assumption 2.2. Consider the rescaled matrix  $A_N(i, j) := \frac{1}{\sqrt{N}} X_N(i, j)$ . Then, the empirical spectral distribution  $\mu_N$  of  $A_N$  converges weakly in probability to the semicircle law  $\nu$  as given by (1.1).

*Remark 2.6.* Conditions (2.3), (2.4), and (2.5) of Assumption 2.3 ensure that the correlations of entries of the stochastic process decay sufficiently fast. A prominent example, in which these conditions are satisfied, is the case where  $(Z_n)_n$  is an ergodic Markov chain on a finite subset of  $\mathbb{R}$  started in its stationary measure  $\rho$ . I.e.  $\mathbb{P}^{Z_1} = \rho$ , and additionally we assume that (2.2) holds true. The verification of (2.3), (2.4), and (2.5) in this case is given in Lemma 3.1 below.

However, there also other interesting processes that satisfy conditions (2.3), (2.4), and (2.5), e.g. certain one-dimensional Gibbs measures on finite state spaces. Moreover, also certain Gaussian processes fall into the realm of Theorem 2.5. These examples will also be treated in the next section.

Regarding the mapping  $\varphi_N$ , we will see in Example 3.6 and Lemma 3.7 that Assumption 2.2 is non-empty. It is e.g. satisfied if we fill the matrix with the entries of the stochastic process one diagonal after another (starting with the main diagonal and always proceeding from top to bottom), i.e.  $\varphi_N(1) = (1, 1)$ ,  $\varphi_N(2) = (2, 2)$ ,  $\dots$ ,  $\varphi_N(N) = (N, N)$ ,  $\varphi_N(N+1) = (1, 2)$ ,  $\varphi_N(N+2) = (2, 3)$ ,  $\dots$ . In view of the results in [9], this is a natural filling. On the other hand, we will also see, that there are fillings for which the convergence from the previous theorem is not true.

*Remark 2.7.* For a stochastic process that satisfies the conditions of Assumption 2.3, we can conclude a further estimate for the left hand side of (2.4): Let  $i_1 \leq \dots \leq i_k$  and  $j_1 \leq \dots \leq j_k$  and let  $l_1 \leq \dots \leq l_{2k}$  be given by  $(l_1, \dots, l_{2k}) := \text{sort}(i_1, \dots, i_k, j_1, \dots, j_k)$ , where  $\text{sort} : \mathbb{R}^{2k} \rightarrow \mathbb{R}^{2k}$  denotes the function that sorts the arguments in increasing order. We claim that for  $L := (l_2 - l_1) + (l_4 - l_3) + \dots + (l_{2k} - l_{2k-1})$  we have

$$\begin{aligned} & |\mathbb{E}(Z_{i_1} \dots Z_{i_k} Z_{j_1} \dots Z_{j_k}) - \mathbb{E}(Z_{i_1} \dots Z_{i_k}) \mathbb{E}(Z_{j_1} \dots Z_{j_k})| \\ & \leq |\mathbb{E}(Z_{i_1} \dots Z_{i_k} Z_{j_1} \dots Z_{j_k})| + |\mathbb{E}(Z_{i_1} \dots Z_{i_k})| \cdot |\mathbb{E}(Z_{j_1} \dots Z_{j_k})| \end{aligned} \quad (2.6)$$

$$\leq C\beta^L. \quad (2.7)$$

The estimate for the first term in (2.6) is obvious from (2.3). For the second term in (2.6) we use again (2.3) for each of the two factors together with

$$\begin{aligned} & (i_2 - i_1) + (i_4 - i_3) + \dots + (i_k - i_{k-1}) \\ & + (j_2 - j_1) + (j_4 - j_3) + \dots + (j_k - j_{k-1}) \geq L. \end{aligned}$$

This can be seen by iterating the simple argument that for any  $a \leq b$  and  $a' \leq b'$  we have

$$(b - a) + (b' - a') \geq \begin{cases} (a' - a) + (b - b') & a \leq a' \leq b' \leq b \\ (a' - a) + (b' - b) & a \leq a' \leq b \leq b'. \end{cases}$$

Combining (2.4) and (2.7) then gives

$$\begin{aligned} |\mathbb{E}(Z_{i_1} \dots Z_{i_k} Z_{j_1} \dots Z_{j_k}) - \mathbb{E}(Z_{i_1} \dots Z_{i_k}) \mathbb{E}(Z_{j_1} \dots Z_{j_k})| & \leq C\beta^{\max(d, L)} \\ & \leq C\beta^{\frac{1}{2}(d+L)}. \end{aligned} \quad (2.8)$$

The number of ways to prove a result like Theorem 2.5 is limited. As it is well known that Wigner's semicircle law  $\nu$  is uniquely determined by its moments, we will use the method of moments, which is one of the fundamental tools when proving limit theorems for dependent random variables. The central step in the proof is thus to show that the moments of the empirical spectral measure converge to the moments of  $\nu$ . These moments are given by the Catalan numbers  $\kappa_k := \frac{1}{k+1} \binom{2k}{k}$ , i.e. (see e.g. [1])

$$\int x^k d\nu(x) = \begin{cases} \kappa_{k/2}, & k \text{ even} \\ 0, & \text{otherwise.} \end{cases}$$

We note that Assumption 2.2 of Theorem 2.5 is sufficient but not necessary. In the following proposition, we provide a condition on the mapping  $\varphi_N$ , under which the empirical eigenvalue distribution does not converge to the semicircle law, if the underlying process is e.g. a certain Markov chain with state space  $\{-1, 1\}$ . This convergence is, of course, already impossible, if the fourth moment of the respective distribution does not converge to  $\kappa_2$ .

**Proposition 2.8.** *Let  $A_N$  be given as in Theorem 2.5. Assume the following:*

- (i) *There exists  $c > 0$  such that  $J(\varphi_N) \geq cN^2$  for all  $N \in \mathbb{N}$ .*
- (ii) *There exist  $\beta \in (0, 1)$ , and  $C > 0$  such that for all  $i_1 \leq i_2 \leq i_3 \leq i_4$  and  $n_j := i_{j+1} - i_j$*

$$\mathbb{E}(Z_{i_1} Z_{i_2} Z_{i_3} Z_{i_4}) = C\beta^{n_1+n_3} \quad (2.9)$$

and

$$\mathbb{E}(Z_{i_1}^2 Z_{i_2}^2) = 1. \quad (2.10)$$

Then we have

$$\lim_{N \rightarrow \infty} \mathbb{E} \left( \frac{1}{N} \text{tr } A_N^4 \right) \neq \kappa_2.$$

Hence, the empirical spectral distribution  $\mu_N$  of  $A_N$  does not converge to  $\nu$ .

*Remark 2.9.* The condition  $J(\varphi_N) \geq cN^2$  of Proposition 2.8 can be interpreted as follows: Considering the path in the upper triangular part of the matrix prescribed by  $\varphi_N$ , if there are ‘too many’ steps from a matrix element to its horizontal or vertical neighbor, the limiting spectral density is not the semicircle.

There are natural examples for fillings with this condition, the most prominent one is probably the mapping

$$\begin{aligned} \varphi_N(1) &= (1, 1), \varphi_N(2) = (1, 2), \dots, \varphi_N(N) = (1, N), \\ \varphi_N(N+1) &= (2, 2), \varphi_N(N+2) = (2, 3), \dots \end{aligned}$$

(i.e. the stochastic process fills the upper triangular matrix row by row from left to right). For this particular mapping we have  $J(\varphi_N) = \frac{N(N+1)}{2} - N + 1 = \frac{N(N-1)}{2} + 1$ .

3. EXAMPLES AND A ‘DIAGONAL’ MAPPING  $\varphi_N$ 

In this section we show that the moment conditions in Assumption 2.3 are satisfied by certain examples, among them Markov chains on a finite state space, high temperature Gibbs measures in dimension one, and Gaussian processes. In Lemma 3.2, we show that there are Markov chains on  $\{1, -1\}$ , that satisfy conditions (2.9) and (2.10). We further provide a mapping  $\varphi_N$  that is valid in the sense of Assumption 2.2. Hence, we show that none of our assumptions in Theorem 2.5 and Proposition 2.8 is empty. We start with the moment conditions in Assumption 2.3.

## 3.1. Markov chains.

**Lemma 3.1.** *Let  $Z_n$  denote a stationary, ergodic Markov chain on a finite subset of  $\mathbb{R}$  with invariant measure  $\rho$ . Assume that  $\mathbb{E}(Z_i^2) = 1$  for all  $i \in \mathbb{N}$  and for all odd numbers  $k$  and all  $i_1, \dots, i_k$  we have that*

$$\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k}) = 0.$$

*Then for all even  $k$  there are constants  $C > 0$  and  $\beta \in [0, 1)$  such that for all  $i_1 \leq i_2 \leq \dots \leq i_k$  it holds*

$$|\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k})| \leq C \beta^{n_1 + n_3 + \dots + n_{k-1}}, \quad n_j := i_{j+1} - i_j \quad (3.11)$$

*and for  $j_1 \leq \dots \leq j_k$  with  $d := \min_{n,m \in \{1, \dots, k\}} |i_n - j_m|$  we have*

$$|\mathbb{E}(Z_{i_1} \dots Z_{i_k} Z_{j_1} \dots Z_{j_k}) - \mathbb{E}(Z_{i_1} \dots Z_{i_k}) \mathbb{E}(Z_{j_1} \dots Z_{j_k})| \leq C \beta^d. \quad (3.12)$$

*Moreover, for  $d' := \min_{j=1,2,\dots} i_{j+1} - i_j$*

$$|\mathbb{E}(Z_{i_1}^2 Z_{i_2}^2 \dots Z_{i_k}^2) - 1| \leq C \beta^{d'}. \quad (3.13)$$

*Proof.* Call the state space of the Markov chain  $S = \{s_1, \dots, s_m\}$ . Note that by standard arguments for any  $l \leq k$

$$\max_{s_1, s_2 \in S} |\mathbb{P}(Z_{i_l} = s_2 | Z_{i_{l-1}} = s_1) - \mathbb{P}(Z_{i_l} = s_2)| \leq C \alpha^{i_l - i_{l-1}}$$

for some constant  $C > 0$  and some  $\alpha \in [0, 1)$  (for a proof see [15], Theorem 4.9). By the Markov property thus also for any  $l \leq k$

$$\begin{aligned} \max_{s_{j_1}, \dots, s_{j_k} \in S} & |\mathbb{P}(Z_{i_k} = s_{j_k}, \dots, Z_{i_l} = s_{j_l} | Z_{i_{l-1}} = s_{j_{l-1}}, \dots, Z_{i_1} = s_{j_1}) \\ & - \mathbb{P}(Z_{i_k} = s_{j_k}, \dots, Z_{i_l} = s_{j_l})| \leq C \alpha^{i_l - i_{l-1}}. \end{aligned}$$

Taking into account the finiteness of the state space  $S$  and that

$$\begin{aligned} \mathbb{E}(Z_{i_k}^{a_{i_k}} \dots Z_{i_l}^{a_{i_l}} | Z_{i_{l-1}} \dots Z_{i_1}) &= \\ \sum_{s_{j_k}, \dots, s_{j_1} \in S} & s_{j_k}^{a_{i_k}} \dots s_{j_l}^{a_{i_l}} \mathbb{1}_{\{Z_{i_{l-1}} = s_{j_{l-1}}, \dots, Z_{i_1} = s_{j_1}\}} \times \\ & \times \mathbb{P}(Z_{i_k} = s_{j_k}, \dots, Z_{i_l} = s_{j_l} | Z_{i_{l-1}} = s_{j_{l-1}}, \dots, Z_{i_1} = s_{j_1}) \end{aligned}$$

for any  $l \leq k$  and any  $i_1 \leq \dots \leq i_{l-1} \leq i_l < i_{l+1} < \dots < i_k$ , we obtain

$$|\mathbb{E}(Z_{i_k}^{a_{i_k}} \dots Z_{i_l}^{a_{i_l}} | Z_{i_{l-1}} \dots Z_{i_1}) - \mathbb{E}(Z_{i_k}^{a_{i_k}} \dots Z_{i_l}^{a_{i_l}})| \leq C \alpha^{i_l - i_{l-1}} \quad (3.14)$$

for all integers  $a_{i_1}, \dots, a_{i_l}$ .

Now, since all odd mixed moments vanish, we obtain for the  $2k$ 'th moments with  $i_1 \leq i_2 \leq \dots \leq i_{2k}$  and all  $l \leq k$

$$\begin{aligned} \mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_{2k}}) &= \mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_{2l-1}} \\ &\quad [\mathbb{E}(Z_{i_{2l}} Z_{i_{2l+1}} \dots Z_{i_{2k}} | Z_{i_1} Z_{i_2} \dots Z_{i_{2l-1}}) - \mathbb{E}(Z_{i_{2l}} \dots Z_{i_{2k}})]). \end{aligned}$$

Thus (3.14) guarantees that

$$|\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_{2k}})| \leq C \alpha^{i_{2l} - i_{2l-1}}$$

for all  $l \leq k$ . Therefore with  $\beta := \alpha^{1/k}$

$$|\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_{2k}})| \leq C \alpha^{\max_{1 \leq l \leq k} \{i_{2l} - i_{2l-1}\}} \leq C \prod_{l=1}^k \beta^{i_{2l} - i_{2l-1}}.$$

This completes the proof of (3.11).

We continue with the proof of (3.13). We observe that (3.14) in particular gives

$$|\mathbb{E}(Z_{i_k}^2 | Z_{i_{k-1}} \dots Z_{i_1}) - \mathbb{E}(Z_{i_k}^2)| = |\mathbb{E}(Z_{i_k}^2 | Z_{i_{k-1}} \dots Z_{i_1}) - 1| \leq C \alpha^{i_k - i_{k-1}}.$$

By the same reasoning as above, together with the basic relation  $\mathbb{E}(XY - 1) = \mathbb{E}(X(Y - 1)) + \mathbb{E}(X - 1)$  (for any random variables  $X$  and  $Y$ ), we have

$$\begin{aligned} |\mathbb{E}(Z_{i_1}^2 Z_{i_2}^2 \dots Z_{i_k}^2) - 1| &= |\mathbb{E}(\mathbb{E}(Z_{i_1}^2 Z_{i_2}^2 \dots Z_{i_k}^2 | Z_{i_{k-1}} \dots Z_{i_1}) - 1)| \\ &= |\mathbb{E}(Z_{i_1}^2 Z_{i_2}^2 \dots Z_{i_{k-1}}^2 \mathbb{E}(Z_{i_k}^2 | Z_{i_{k-1}} \dots Z_{i_1}) - 1)| \\ &= |\mathbb{E}(Z_{i_1}^2 Z_{i_2}^2 \dots Z_{i_{k-1}}^2 [\mathbb{E}(Z_{i_k}^2 | Z_{i_{k-1}} \dots Z_{i_1}) - 1]) + \mathbb{E}(Z_{i_1}^2 Z_{i_2}^2 \dots Z_{i_{k-1}}^2 - 1)| \\ &\leq C \alpha^{i_k - i_{k-1}} + |\mathbb{E}(Z_{i_1}^2 Z_{i_2}^2 \dots Z_{i_{k-1}}^2 - 1)|. \end{aligned}$$

Iterating this calculation shows (3.13).

We continue with the proof of (3.12). We denote by  $l_1 \leq \dots \leq l_{2k}$  the (sorted) joint indices consisting of  $i_1, \dots, i_k$  and  $j_1, \dots, j_k$ , i.e.  $(l_1, \dots, l_{2k}) = \text{sort}(i_1, \dots, i_k, j_1, \dots, j_k)$ , where  $\text{sort} : \mathbb{R}^{2k} \rightarrow \mathbb{R}^{2k}$  denotes the function that sorts the arguments in increasing order. We have the general identity for arbitrary  $1 \leq n \leq 2k - 1$

$$\begin{aligned} &|\mathbb{E}(Z_{l_1} \dots Z_{l_{2k}}) - \mathbb{E}(Z_{l_1} \dots Z_{l_n}) \mathbb{E}(Z_{l_{n+1}} \dots Z_{l_{2k}})| \\ &= |\mathbb{E}(\mathbb{E}(Z_{l_1} \dots Z_{l_{2k}} | Z_{l_1} \dots Z_{l_n}) - Z_{l_1} \dots Z_{l_n} \mathbb{E}(Z_{l_{n+1}} \dots Z_{l_{2k}}))| \\ &= |\mathbb{E}(Z_{l_1} \dots Z_{l_n} [\mathbb{E}(Z_{l_{n+1}} \dots Z_{l_{2k}} | Z_{l_1} \dots Z_{l_n}) - \mathbb{E}(Z_{l_{n+1}} \dots Z_{l_{2k}})])| \\ &\leq C \alpha^{l_{n+1} - l_n}, \end{aligned} \tag{3.15}$$

where the last estimate is due to (3.14). If  $i_k \leq j_1$  or  $j_k \leq i_1$  this proves the claim. To complete the proof for an arbitrary ordering of the indices, we apply the above estimate successively for certain values of  $n$ . Without loss of generality, we assume that  $l_1 = i_1$ . Let  $n_1, n_2, \dots, n_m$  denote the indices such



that

$$\begin{aligned}(l_1, \dots, l_{n_1}) &= (i_1, \dots, i_{n_1}) \\ (l_{n_1+1}, \dots, l_{n_2}) &= (j_1, \dots, j_{n_2}) \\ (l_{n_2+1}, \dots, l_{n_3}) &= (i_{n_1+1}, \dots, i_{n_1+(n_3-n_2)}) \\ (l_{n_3+1}, \dots, l_{n_4}) &= (j_{n_2+1}, \dots, j_{n_2+(n_4-n_3)}), \dots\end{aligned}$$

This means  $(l_1, \dots, l_{n_1}), (l_{n_1+1}, \dots, l_{n_2}), \dots$  are the longest subsequences of  $(l_1, \dots, l_{2k})$  such that each sub-sequence consists either of indices in  $\{i_1, \dots, i_k\}$  or in  $\{j_1, \dots, j_k\}$  only. Hence, on the one hand, by iterating (3.15), we obtain for some (possibly different) constant  $C$

$$\begin{aligned}& |\mathbb{E}(Z_{l_1} \dots Z_{l_{2k}}) - \mathbb{E}(Z_{l_1} \dots Z_{l_{n_1}}) \mathbb{E}(Z_{l_{n_1+1}} \dots Z_{l_{n_2}}) \dots \mathbb{E}(Z_{l_{n_{m-1}+1}} \dots Z_{l_{n_m}})| \\ & \leq C \alpha^d.\end{aligned}\tag{3.16}$$

On the other hand, we can apply the analogous factorization to both terms  $\mathbb{E}(Z_{i_1} \dots Z_{i_k})$  and  $\mathbb{E}(Z_{j_1} \dots Z_{j_k})$  and use e.g.  $i_{n_1+1} - i_{n_1} \geq d, j_{n_2+1} - i_{n_2} \geq d, \dots$  to obtain

$$\begin{aligned}& |\mathbb{E}(Z_{i_1} \dots Z_{i_k}) \mathbb{E}(Z_{j_1} \dots Z_{j_k}) \\ & - \mathbb{E}(Z_{l_1} \dots Z_{l_{n_1}}) \mathbb{E}(Z_{l_{n_1+1}} \dots Z_{l_{n_2}}) \dots \mathbb{E}(Z_{l_{n_{m-1}+1}} \dots Z_{l_{n_m}})| \leq C \alpha^d.\end{aligned}\tag{3.17}$$

Combining (3.16) and (3.17) completes the proof of (3.12).  $\square$

In order to see that the assumption of Proposition 2.8 is non-empty, we consider Markov chains with state space  $\{-1, +1\}$  and  $\mathbb{P}(Z_n = i | Z_{n-1} = i) = p$ . As in this case trivially  $\mathbb{E}(Z_{i_1}^2 Z_{i_2}^2) = 1$ , it remains to verify (2.9). Lemma 3.2 shows for  $p > \frac{1}{2}$  that (2.9) is valid with  $C = 1$  and  $\beta = 2p - 1$ . Hence, such a Markov chain together with a filling that satisfies  $J(\varphi_N) \geq cN^2$ , e.g. a row-wise filling, leads to a limiting spectral measure that differs from the semi-circle (see Fig. 1). In addition, such a Markov chain also falls in the realm of Lemma 3.1 and hence together with a filling that satisfies Assumption 2.2, generates random matrices, for which the limiting spectral measure is the semicircle. Thus, in this case, the filling  $\varphi_N$  is essential for the limiting spectral measure.

**Lemma 3.2.** *Let  $Z_n$  denote a Markov chain with state space  $S = \{-1, +1\}$  and transition matrix*

$$\begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix}$$

*for some  $p \in (0, 1)$ . Let  $\mathbb{P}(Z_1 = 1) = \mathbb{P}(Z_1 = -1) = \frac{1}{2}$ . Further set  $\beta := 2p - 1$ . Then we have for  $k \in \mathbb{N}, i_1 \leq i_2 \leq \dots \leq i_k$  and  $n_j := i_{j+1} - i_j$*

$$\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k}) = \begin{cases} \beta^{n_1+n_3+\dots+n_{k-1}}, & k \text{ even} \\ 0, & \text{else} \end{cases}.$$

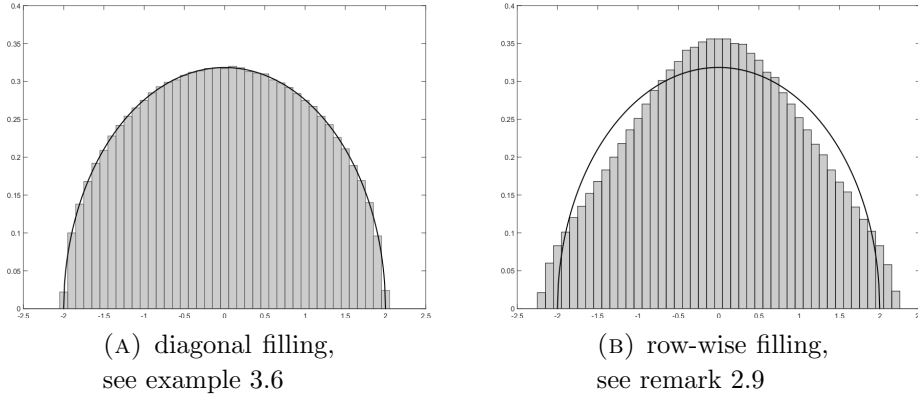


FIGURE 1. Histograms of eigenvalues of a  $10000 \times 10000$  matrix, filled with a Markov chain  $(Z_n)_n$  with state space  $\{-1, 1\}$  and  $\mathbb{P}(Z_n = i | Z_{n-1} = i) = 0.7$ . The line shows the density of the semi-circle.

*Proof.* The trick is similar to the previous proof. However, for a binary Markov chain we can compute explicitly that

$$\mathbb{E}(Z_{i+j} | Z_i) = (2p - 1)^j Z_i.$$

Therefore, using that  $s^2 = 1$  for all  $s \in S$  we obtain for  $i_1 < i_2 < \dots < i_k$

$$\begin{aligned} \mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k}) &= \mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_{k-1}} \mathbb{E}(Z_{i_k} | Z_{i_{k-1}})) \\ &= (2p - 1)^{i_k - i_{k-1}} \mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_{k-2}} Z_{i_{k-1}}^2) \\ &= (2p - 1)^{i_k - i_{k-1}} \mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_{k-2}}). \end{aligned}$$

Repeating this calculation successively, we obtain

$$\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k}) = \begin{cases} \beta^{n_1 + n_3 + \dots + n_{k-1}}, & k \text{ even} \\ 0, & \text{else} \end{cases}$$

as asserted. □

**3.2. Gibbs measures.** The second important class of examples of random variables that satisfy Assumption 2.3 is provided by some Gibbs measures in one dimension. We prove

**Lemma 3.3.** *Let  $Z_n$  denote a stochastic process given by a one-dimensional Gibbs measure for some shift-invariant potential*

$$\Phi = \{\phi_A : A \subset \mathbb{Z}, 0 < |A| < \infty\}$$

*on a finite state space (here we are following the Georgii's definition of a Gibbs measure in [10]). Assume that for all odd numbers  $k$  and all  $i_1, \dots, i_k$  we have that*

$$\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k}) = 0$$

and  $\mathbb{E}(Z_i^2) = 1$  for all  $i \in \mathbb{N}$ . Moreover, assume that Dobrushin's condition is satisfied, i.e.

$$\sum_{A:0 \in A} (|A| - 1) \sup_{\zeta, \eta} |\phi_A(\zeta) - \phi_A(\eta)| < 2. \quad (3.18)$$

holds.

Finally we require the following condition

$$\sum_{A:0 \in A} e^{t \text{diam}(A)} (|A| - 1) \sup_{\zeta, \eta} |\phi_A(\zeta) - \phi_A(\eta)| < \infty, \quad (3.19)$$

for some  $t > 0$ .

Then for all even  $k$  there are constants  $C > 0$  and  $\beta \in [0, 1)$  such that for all  $i_1 \leq i_2 \leq \dots \leq i_k$  with  $n_j := i_{j+1} - i_j$  it holds

$$|\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k})| \leq C \beta^{n_1 + n_3 + \dots + n_{k-1}}$$

and for  $j_1 \leq \dots \leq j_k$  with  $d := \min_{n,m \in \{1, \dots, k\}} |i_n - j_m|$  we have

$$|\mathbb{E}(Z_{i_1} \dots Z_{i_k} Z_{j_1} \dots Z_{j_k}) - \mathbb{E}(Z_{i_1} \dots Z_{i_k}) \mathbb{E}(Z_{j_1} \dots Z_{j_k})| \leq C \beta^d.$$

Moreover, for  $d' := \min_{j=1,3,5,\dots} i_{j+2} - i_j$  it holds

$$\left| \mathbb{E}(Z_{i_1}^2 Z_{i_3}^2 \dots Z_{i_{k-1}}^2) - 1 \right| \leq C \beta^{d'}.$$

*Proof.* Note that the crucial estimate in the proof of Lemma 3.1 was that for any  $l \leq k$  and all  $i_1 \leq i_2 \leq \dots \leq i_k$  we have

$$\begin{aligned} \max_{s_{j_1}, \dots, s_{j_k} \in S} |\mathbb{P}(Z_{i_k} = s_{j_k}, \dots, Z_{i_l} = s_{j_l} | Z_{i_{l-1}} = s_{j_{l-1}}, \dots, Z_{i_1} = s_{j_1}) \\ - \mathbb{P}(Z_{i_k} = s_{j_k}, \dots, Z_{i_l} = s_{j_l})| \leq C \alpha^{i_l - i_{l-1}} \end{aligned}$$

for some constant  $C > 0$  and some  $\alpha \in [0, 1)$ . Now due to [10], Chapter 8, this estimate holds if conditions (3.18) and (3.19) are satisfied. Thus the rest of the proof follows the proof of Lemma 3.1.  $\square$

*Remark 3.4.* If we take, for example, a potential with finite range, then (3.19) is satisfied. If we consider a Gibbs measure with a parameter  $\beta > 0$  (usually called the inverse temperature), i.e. we substitute  $\Phi$  by  $\Phi_\beta := \{\beta \phi_A : A \subset \mathbb{Z}, 0 < |A| < \infty\}$ , then (3.18) holds whenever  $\beta$  is small enough.

**3.3. Gaussian processes.** Also Gaussian Markov processes satisfy the assumptions of Theorem 2.5. More precisely we prove:

**Lemma 3.5.** *Let  $(Z_n)$  denote a stationary Gaussian Markov process with zero mean and variance one. For  $k \in \mathbb{N}$  there are constants  $C > 0$  and  $\beta \in [0, 1)$  such that for all  $i_1 \leq i_2 \leq \dots \leq i_k$  with  $n_j := i_{j+1} - i_j$  and  $j_1 \leq \dots \leq j_k$  with  $d := \min_{n,m \in \{1, \dots, k\}} |i_n - j_m|$  it holds*

$$|\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k})| \leq C \beta^{n_1 + n_3 + \dots + n_{k-1}}, \quad \text{if } k \text{ even}, \quad (3.20)$$

$$|\mathbb{E}(Z_{i_1} Z_{i_2} \dots Z_{i_k})| = 0, \quad \text{if } k \text{ odd} \quad (3.21)$$

and

$$|\mathbb{E}(Z_{i_1} \dots Z_{i_k} Z_{j_1} \dots Z_{j_k}) - \mathbb{E}(Z_{i_1} \dots Z_{i_k}) \mathbb{E}(Z_{j_1} \dots Z_{j_k})| \leq C \beta^d. \quad (3.22)$$

Moreover, for  $d' := \min_{j=1,3,5,\dots} i_{j+2} - i_j$  we have

$$\left| \mathbb{E}(Z_{i_1}^2 Z_{i_3}^2 \dots Z_{i_{k-1}}^2) - 1 \right| \leq C\beta^{d'}. \quad (3.23)$$

*Proof.* First, observe that assertion (3.21) is obvious. Assertions (3.20), (3.22) and (3.23) follow from Isserli's theorem together with Doob's theorem. Indeed, in view of (3.20), Isserlis' theorem [14] states that for all even  $k$  and all indices  $i_1 \leq i_2 \leq \dots \leq i_{2k}$

$$\mathbb{E}[Z_{i_1} \dots Z_{i_{2k}}] = \prod_{l=1}^k t(i_{2l-1}, i_{2l}) + \sum_{\sigma \in S_{2k}^k} \prod_{l=1}^k t(i_{\sigma(2l-1)}, i_{\sigma(2l)}), \quad (3.24)$$

where we denote by  $S_{2k}^k$  the set of all permutations  $\sigma$  of  $\{1, \dots, 2k\}$  such that  $\{\sigma(2l-1), \sigma(2l)\} \neq \{2l'-1, 2l'\}$  for at least one  $l \in \{1, \dots, k\}$  and any  $l' \in \{1, \dots, k\}$  and for any  $i, i' \in \mathbb{N}$ ,

$$t(i, i') := \text{Cov}(Z_i, Z_{i'}),$$

are the covariances.

As a consequence of Doob's theorem [4] we can conclude that for Gaussian Markov process  $t(i, i') = \beta^{|i-i'|}$  for some  $-1 < \beta < 1$  and any  $i, i' \in \mathbb{N}$ . Thus the summands in the second term on the right hand side of (3.24) are all smaller than the first summand. Together with the fact, that the size of  $S_{2k}^k$  is finite and only depends on  $k$  this shows (3.20). By similar arguments, we obtain (3.22) and (3.23).  $\square$

**3.4. A 'diagonal' mapping  $\varphi_N$ .** In the following example and lemma, we show that there are fillings of a random matrix by a stochastic process that satisfy Assumption 2.2. Fortunately, the situation we always had in mind, namely the matrix we obtain when writing the entries of the stochastic process successively on the diagonals of the matrix as indicated by remark 2.9 satisfies Assumption 2.2.

*Example 3.6.* We consider the mapping  $\varphi_N$  that leads to the following matrices:

$$X_N = \begin{pmatrix} Z_1 & Z_{N+1} & Z_{2N} & \dots & \dots & Z_{\frac{N(N+1)}{2}} \\ Z_{N+1} & Z_2 & Z_{N+2} & \dots & \dots & \dots \\ Z_{2N} & Z_{N+2} & Z_3 & Z_{N+3} & \dots & \dots \\ \dots & \dots & Z_{N+3} & Z_4 & \dots & \dots \\ \dots & \dots & \dots & Z_{10} & \dots & Z_{2N-1} \\ Z_{\frac{N(N+1)}{2}} & \dots & \dots & \dots & Z_{2N-1} & Z_N \end{pmatrix}.$$

To check whether Assumption 2.2 is satisfied, we need to compute the distance defined in Definition 2.1, (b). The main ingredient to calculate this distance is a representation of the inverse  $\varphi_N^{(-1)}$ . In this case, it can be calculated via

$$\varphi_N^{(-1)}(j, k) = \frac{N(N+1)}{2} - \frac{(N-|j-k|)(N-|j-k|+1)}{2} + \min(j, k).$$

For  $j \leq k$  this formula can easily be verified as follows: The first term  $\frac{N(N+1)}{2}$  denotes the number of steps between the  $(1, 1)$ -entry and the  $(1, N)$ -entry. Observe that  $N - (k - j)$  is the length of the diagonal containing the  $(j, k)$ -entry. Hence, the difference of the first two terms is the number of steps from the  $(1, 1)$ -entry to the upper left corner of the diagonal with the  $(j, k)$ -entry. Further there are  $j$  steps along this diagonal to the  $(j, k)$  entry.

We will now verify that this filling is indeed valid in the sense of Assumption 2.2.

**Lemma 3.7.** *For  $\varphi_N$  as in Example 3.6, we have for all  $N \in \mathbb{N}$ ,  $i, j \in \{1, \dots, N\}$  and  $n \in \mathbb{N}, n > 0$*

$$\#\{x \in \{1, \dots, N\} : \|(i, x) - (x, j)\|_{\varphi_N} = n\} \leq 4.$$

*In particular, Assumption 2.2 is satisfied and for stochastic processes that satisfy Assumption 2.3 the limiting spectral density is given by the semi-circle.*

*Proof.* Let  $N, i, j$  be fixed. We assume without loss of generality that  $i < j$  (observe that  $n \neq 0$  implies  $i \neq j$ ). We consider

$$F(x) := \|(i, x) - (x, j)\|_{\varphi_N}.$$

The main idea is that  $\{1, \dots, N\}$  can be divided into four subsets, where  $F$  is strictly monotone on each of the subsets. Hence, for given  $n$ ,  $F(x) = n$  for at most four values of  $x$ . We consider these subsets individually and recall the representation of  $\varphi_N^{(-1)}$  given in example 3.6, i.e.

$$\begin{aligned} F(x) &= \left| \frac{(N - |i - x|)(N - |i - x| + 1) - (N - |x - j|)(N - |x - j| + 1)}{2} \right. \\ &\quad \left. + \min(i, x) - \min(x, j) \right| \\ &= \left| \frac{i^2 - j^2 + 2x(j - i) - (2N + 1)[|i - x| - |j - x|]}{2} \right. \\ &\quad \left. + \min(i, x) - \min(x, j) \right|. \end{aligned}$$

Case 1:  $x \leq i$  We will show that  $F(x)_{x \leq i}$  is strictly increasing. We have

$$F(x) = (j - i) \left| -\frac{1}{2}(i + j) + x + N + \frac{1}{2} \right| = x(j - i) + C(i, j, N)$$

for some constant  $C(i, j, N)$ , which does not depend on  $x$ . Hence,  $F(x)$  is obviously strictly increasing for  $x \leq i$ .

Case 2:  $x \geq j$  We have

$$F(x) = (j - i) \left| x - \frac{1}{2}(j + i) - N - \frac{3}{2} \right| = -x(j - i) + C(i, j, N).$$

Here, again  $C(i, j, N)$  denotes a constant, which does not depend on  $x$  (though it may differ from case 1). Hence,  $F(x)$  is strictly decreasing in this case.

Case 3:  $i < x < j$  We have by an easy calculation

$$F(x) = \left| ((j-i) - (2N+1)) \left( x - \frac{1}{2}(i+j) \right) + i - x \right|$$

On the one hand, as  $(j-i) - (2N+1) \leq 0$  and  $i - x \leq 0$ , the term within the absolute value is negative for  $x \geq \frac{1}{2}(i+j)$ . On the other hand, the term is positive for  $x < \frac{1}{2}(i+j)$  (this follows from  $(2N+1) - (j-i) \geq x - i$ ). Hence we have for some constant  $C(i, j, N)$

$$F(x) = \begin{cases} ((j-i) - (2N+2))x - C(i, j, N), & x < \frac{1}{2}(i+j) \\ -((j-i) - (2N+2))x + C(i, j, N), & x \geq \frac{1}{2}(i+j) \end{cases},$$

which shows that  $F(x)$  is decreasing for  $i < x < \frac{1}{2}(i+j)$  and increasing for  $\frac{1}{2}(i+j) \leq x < j$ .  $\square$

#### 4. PROOF OF THE MAIN RESULT

Before we complete the proof of the main result, we introduce some notation. For  $k \in \mathbb{N}$  we set

$$S_k := \{P = (P_1, \dots, P_k) : P_i = (p_i, p_{i+1}) \in \{1, \dots, N\}^2\}.$$

Then we can write (with  $k+1 = 1$ )

$$\begin{aligned} \mathbb{E} \left( \frac{1}{N} \text{tr} A_N^k \right) &= \frac{1}{N^{k/2+1}} \sum_{P \in S_k} \mathbb{E} [X_N(P_1) X_N(P_2) \dots X_N(P_k)] \\ &= \frac{1}{N^{k/2+1}} \sum_{P \in S_k} \mathbb{E} [X_N(P)], \end{aligned} \quad (4.25)$$

where we further abbreviated  $X_N(P) := X_N(P_1) X_N(P_2) \dots X_N(P_k)$ . In order to prove Theorem 2.5 by the method of moments, it suffices to prove the following lemma (see [1, Lemma 2.1.6 and Lemma 2.1.7]).

**Lemma 4.1.** *For  $X_N \in \mathbb{R}^{N \times N}$  generated by a stochastic process  $Z_n$  and a mapping  $\varphi_N$  as in Theorem 2.5 and  $A_N(i, j) := \frac{1}{\sqrt{N}} X_N(i, j)$  we have*

(i)

$$\lim_{N \rightarrow \infty} \mathbb{E} \left( \frac{1}{N} \text{tr} A_N^k \right) = \begin{cases} 0 & k \text{ odd} \\ \kappa_{k/2}, & k \text{ even} \end{cases}.$$

(ii)

$$\lim_{N \rightarrow \infty} \frac{1}{N^{k+2}} \sum_{P \in S_k, Q \in S_k} \mathbb{E} [X_N(P) X_N(Q)] - \mathbb{E} [X_N(P)] \mathbb{E} [X_N(Q)] = 0$$

*Proof.* We start with the proof of (i) and consider the representation given in (4.25). By (2.2), we can assume that  $k$  is even, as the term in (4.25) vanishes for  $k$  odd. For  $P \in S_k$  we further introduce the notation

$$[i]_P := \{j \neq i : P_i = P_j \text{ or } P_i = (p_{j+1}, p_j)\}$$

and we write

- $P \in S_k^0$ , if  $\#[i]_P = 2$  for all  $i = 1, \dots, k$
- $P \in S_k^1$ , if  $\#[i]_P$  even for all  $i = 1, \dots, k$  and there exists  $i_0$  such that  $\#[i_0]_P \geq 4$
- $P \in S_k^*$ , if there exists  $i_0$  such that  $\#[i_0]_P$  is odd.

We observe that by a simple combinatorial argument, we have

$$\#\{P : P \in S_k^1\} = o(N^{k/2+1}). \quad (4.26)$$

Indeed, assume that an equivalence relation on the indices  $1, \dots, k$  with equivalence classes  $[i]$  is given such that each equivalence class has an even number of elements and one equivalence class has at least four elements (the amount of such equivalence relations depends on  $k$  only). Now, we estimate the number of  $P$ , such that  $[i] = [i]_P$  for all  $i$ . There are at most  $N^2$  choices for  $(p_1, p_2)$ . When determining the values  $p_3, p_4, \dots$  successively, there are at most  $N$  choices for each  $p_{j+1}$  where  $j \notin [i]$  for  $i < j$  (otherwise there are at most 2 possibilities). Observe, that there are at most  $\frac{k}{2} - 1$  equivalence classes in order to ensure  $P \in S_k^1$ . Hence, we can neglect those terms with  $P \in S_k^1$  in (4.25) and we have

$$\mathbb{E} \left( \frac{1}{N} \text{tr} A_N^k \right) = \frac{1}{N^{k/2+1}} \sum_{P \in S_k^0 \cup S_k^*} \mathbb{E} [X_N(P_1) X_N(P_2) \dots X_N(P_k)] + o(1). \quad (4.27)$$

Next, we argue that the terms with  $P \in S_k^0$  give a contribution of  $\kappa_{k/2}$  in (4.27), i.e. we show

$$\frac{1}{N^{k/2+1}} \sum_{P \in S_k^0} \mathbb{E} [X_N(P_1) X_N(P_2) \dots X_N(P_k)] = \kappa_{k/2} + o(1). \quad (4.28)$$

For  $k = 2$  the proof of (4.28) is immediate from

$$\frac{1}{N^2} \sum_{P \in S_2^0} \mathbb{E} [X_N(P_1)^2] = 1 = \kappa_1.$$

The main ingredient for the proof of (4.28) for  $k \geq 4$  is (2.5), i.e. for  $i_1 < i_3 < \dots < i_{k-1}$  with  $d := \min_{j=1,3,5,\dots} i_{j+2} - i_j$  we have

$$\mathbb{E}(Z_{i_1}^2 Z_{i_3}^2 \dots Z_{i_{k-1}}^2) = 1 + \mathcal{O}(\beta^d). \quad (4.29)$$

Here, the constant implicit in the  $\mathcal{O}$ -term does not depend on  $i_1, i_3, \dots, i_{k-1}$ . In order to use (4.29), we need to introduce a sorting procedure that allows us to write  $X_N(P_1) \dots X_N(P_k)$  in terms of the stochastic process  $(Z_n)_n$  in increasing order of the indices. We set

$$G_N : (\{1, \dots, N\}^2)^k \rightarrow \left\{ 1, \dots, \frac{N(N+1)}{2} \right\}^k;$$

$$G_N(P) := \text{sort}(\varphi_N^{(-1)}(P_1), \varphi_N^{(-1)}(P_2), \dots, \varphi_N^{(-1)}(P_k))$$

Here,  $\text{sort}(\cdot) : \mathbb{R}^k \rightarrow \mathbb{R}^k$  denotes the function that permutes the arguments such that they are in increasing order. Hence, for  $P \in S_k$  with  $G_N(P) = (i_1, \dots, i_k)$ ,

we have

$$X_N(P_1)X_N(P_2)\dots X_N(P_k) = Z_{i_1}Z_{i_2}\dots Z_{i_k}, \quad i_1 \leq \dots \leq i_k. \quad (4.30)$$

In view of (4.29), we set for  $P \in S_k^0$  with  $G_N(P) = (i_1, i_1, \dots, i_{k-1}, i_{k-1})$

$$d(P) := \min_{j=1,3,5,\dots} (i_{j+2} - i_j).$$

With this notation and (4.29) we obtain

$$\frac{1}{N^{k/2+1}} \sum_{P \in S_k^0} \mathbb{E}[X_N(P)] = \frac{\#S_k^0}{N^{k/2+1}} + \frac{1}{N^{k/2+1}} \mathcal{O} \left( \sum_{P \in S_k^0} \beta^{d(P)} \right),$$

where the constant implicit in the  $\mathcal{O}$ -term is independent of  $N$ . By the same combinatorial arguments like in the classical proof of Wigner's semicircle law by the moment method (see [2] or [1]) we have

$$\lim_{N \rightarrow \infty} \frac{\#S_k^0}{N^{k/2+1}} = \kappa_{k/2}.$$

Hence, to show (4.28), it remains to prove

$$\frac{1}{N^{k/2+1}} \sum_{P \in S_k^0} \beta^{d(P)} = o(1). \quad (4.31)$$

We distinguish between  $P \in S_k^0$  with  $d(P) \geq \sqrt{N}$  and with  $d(P) < \sqrt{N}$ . For  $d(P) \geq \sqrt{N}$ , we use  $\beta^{d(P)} \leq \beta^{\sqrt{N}}$  and by  $\#S_k^0 = \mathcal{O}(N^{k/2+1})$ , we have

$$\frac{1}{N^{k/2+1}} \sum_{P \in S_k^0, d(P) \geq \sqrt{N}} \beta^{d(P)} = o(1).$$

For the remaining terms with  $d(P) \leq \sqrt{N}$  we estimate  $\beta^{d(P)} \leq 1$  and hence for (4.31) it is sufficient to prove  $\#\{P \in S_k^0 : d(P) < \sqrt{N}\} = o(N^{k/2+1})$ . We can proceed very similarly as in the proof of (4.26). Neglecting the additional information  $d(P) < \sqrt{N}$  for a moment, we consider  $k/2$  equivalence classes  $[i]$  and determining  $p_1, p_2, \dots$  successively, there are  $N$  choices for each 'new' equivalence class and  $N$  additionally choices for the 'starting point', say  $p_1$ . This would lead to an upper bound of  $N^{k/2+1}$ , which is not sufficient. However, the restriction  $d(P) < \sqrt{N}$  means that there are two equivalence classes  $[i]$  and  $[j]$ ,  $i < j$  (again the number of such possible pairs of equivalence classes depends on  $k$  only) such that  $\|P_i - P_j\|_{\varphi_N} \leq \sqrt{N}$ . Hence, in our procedure, when we already determined  $P_i$  and first encounter the equivalence class  $[j]$ , we actually have only  $2\sqrt{N}$  choices rather than  $N$ . This reduces the bound for the total number of possibilities to  $N^{(k+1)/2}$ , which shows

$$\#\{P \in S_k^0 : d(P) < \sqrt{N}\} = o(N^{k/2+1})$$

and hence completes the proof of (4.31) resp. of (4.28). So far, we have

$$\mathbb{E} \left( \frac{1}{N} \text{tr} A_N^k \right) = \kappa_{k/2} + \frac{1}{N^{k/2+1}} \sum_{P \in \cup S_k^*} \mathbb{E}[X_N(P_1)X_N(P_2)\dots X_N(P_k)] + o(1) \quad (4.32)$$



and it remains to show that the second term on the r.h.s. vanishes as  $N \rightarrow \infty$ .

When sorting  $X_N(P_1)X_N(P_2)\dots$  as in (4.30), we recall that by (2.3) the differences  $i_2 - i_1, i_4 - i_3, \dots$  are of particular interest. We model this structure with the help of pair partitions. Note that, in the following paragraphs we will use partitions in a slightly different way than in the above paragraph. Recall that in the derivation of (4.32),  $i \sim j$  corresponded to  $(p_i, p_{i+1}) = (p_j, p_{j+1})$  or  $(p_i, p_{i+1}) = (p_{j+1}, p_j)$ . From now on, we use  $i \sim j$  to model that after sorting according to (4.30),  $X_N(P_i)$  and  $X_N(P_j)$  are neighbors. We introduce

$$\mathcal{PP}(k) := \{\pi : \pi \text{ is a pair partition of } \{1, \dots, k\}\}.$$

For  $\pi \in \mathcal{PP}(k)$ , we say that  $P \in S_k$  with  $(i_1, \dots, i_k) = G_N(P)$  (see (4.30)) is  $\pi$ -consistent, if

$$i \sim_\pi j \Leftrightarrow \{\varphi_N^{(-1)}(P_i), \varphi_N^{(-1)}(P_j)\} = \{i_l, i_{l+1}\} \quad \text{for some odd } l.$$

The  $\pi$ -consistency of  $P$ , we write  $P \in S_k(\pi)$ , means that  $\pi$  prescribes the pairs  $P_i, P_j$  that correspond to an odd and the preceding even position of the vector obtained after applying  $G_N$  to  $P$ .

We further introduce the notation for  $n_1, \dots, n_k \in \mathbb{N}$  and  $\pi \in \mathcal{PP}(k)$

$$\begin{aligned} m(n_1, n_3, \dots, n_{k-1}, \pi) \\ := \#\{P \in S_k(\pi) : (G_N(P))_{j+1} - (G_N(P))_j = n_j, j = 1, 3, 5, \dots\}. \end{aligned}$$

Then  $m$  is the number of possible vectors  $P$ , consistent with  $\pi$ , such that, after sorting  $X_N(P_1)X_N(P_2)\dots X_N(P_k)$  in the order prescribed by the underlying stochastic process, the first and the second terms are separated by  $n_1$  steps of the stochastic process, the third and the fourth term are separated by  $n_3$  steps in the stochastic process and so on. We obtain by (2.3) that there are constants  $C > 0$  and  $\beta \in [0, 1)$  such that

$$\begin{aligned} & \left| \frac{1}{N^{k/2+1}} \sum_{P \in S_k^*} \mathbb{E}[X_N(P_1)X_N(P_2)\dots X_N(P_k)] \right| \\ & \leq C \frac{1}{N^{k/2+1}} \sum_{\pi \in \mathcal{PP}(k)} \sum_{\substack{n_1, n_3, \dots \geq 0, \\ \exists i: n_i \neq 0}} m(n_1, n_3, \dots, n_{k-1}, \pi) \beta^{n_1+n_3+\dots+n_{k-1}}. \end{aligned}$$

Recalling that  $0 \leq \beta < 1$ , we have  $\sum_{n_1, n_3, \dots \geq 0} \beta^{n_1+n_3+\dots+n_{k-1}} \leq K$  for some constant  $K$ . Hence it suffices to show that for any  $n_1, n_3, \dots \geq 0$  with some  $n_i > 0$  and any pair partition  $\pi$ , we have

$$m(n_1, \dots, n_{k-1}, \pi) = o(N^{k/2+1}). \quad (4.33)$$

To prove (4.33), let  $n_1, n_3, \dots \geq 0$  be fixed and without loss of generality we assume that  $n_1 > 0$ . We distinguish between crossing partitions  $\pi$  and non-crossing partitions. Here, a partition is said to be crossing if there are indices  $i, i', j, j'$  with

$$i \sim_\pi j, \quad i' \sim_\pi j' \quad \text{and} \quad i < i' < j < j'. \quad (4.34)$$

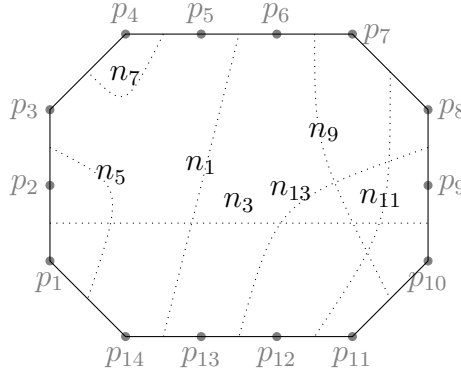


FIGURE 2. Example of a crossing partition with  $k = 14$ ; dotted lines indicate the pair partition.

First, we show that for all  $n_1, \dots, n_{k-1}$  and all crossing pair partitions  $\pi$  we have

$$m(n_1, \dots, n_{k-1}, \pi) \leq CN^{k/2} \quad (4.35)$$

by estimating the number of possible choices for  $P$  that contribute to the number  $m(n_1, \dots, n_{k-1}, \pi)$ . We note that the number of possibilities to sort the  $\frac{k}{2}$  partition blocks depends on  $k$  only, and we may hence assume that each partition block is associated to one of the  $n_1, n_3, \dots$ . Since  $\pi$  is a crossing partitions there are indices  $i, i', j, j'$  with (4.34). Without loss of generality we can assume that  $i = 1$ . An example for  $k = 14, i = 1, j = 9, i' = 2, j' = 14$  is shown in Fig. 2.

The main idea is to choose  $p_1, p_2, \dots$  successively (in an appropriate order), where once we fixed  $P_i = (p_i, p_{i+1})$  there are only 4 possible choices for  $P_j$  if  $i \sim j$  (as we already decided which value of  $n_1, n_3, \dots$  is associated to this pair, say  $n_1$ , we either have to choose  $P_j$  such that it is  $n_1$  steps before or  $n_1$  steps after  $P_i$  in the stochastic process, leading to 2 possibilities; considering the symmetry of the matrix we have a total of 4 possibilities for  $P_j$  if  $P_i$  is fixed). We proceed in the following order: We chose  $P_1 = (p_1, p_2)$ , for which there are at most  $N^2$  possibilities. We further choose  $P_2, P_3, \dots, P_{i-1} = (p_{i-1}, p_i)$  with  $N$  possibilities whenever we encounter a ‘new’ equivalence class and 4 possibilities otherwise. As  $1 \sim j$ , we have only 4 possibilities for  $P_j$ . Now, we proceed further with  $P_{j-1}, P_{j-2}, \dots, P_{i+1}$  in the same way as before. Then we already determined  $p_i$  and  $p_{i+1}$  (and hence  $P_i$ ). We continue in the same way with the remaining points. Hence, we had  $N$  possibilities for  $p_1$  and  $N$  possibilities for each of the  $\frac{k}{2}$  equivalence classes, except for the equivalence class containing  $i$ . Hence, we had  $CN^{\frac{k}{2}}$  possible choices, which proves (4.35).

It remains to show that

$$m(n_1, \dots, n_{k-1}, \pi) = o(N^{k/2+1}), \quad \text{if } \pi \in \mathcal{PP}(k) \text{ is non-crossing.} \quad (4.36)$$

We recall that  $n_1 \neq 0$ . We observe that each non-crossing partition has a block of the form  $\{i, i+1\}$ , since for any  $i \sim j, j > i+1$  there exist  $i < i' < j' < j$

with  $i' \sim j'$ . If the associated  $n_i$  value is zero, we have  $p_i = p_{j+1}$  and hence we can first determine the sequence  $P_1, P_2, \dots, P_{i-1}, P_{i+2}, \dots$  and obtain the original sequence by further  $N$  choices for  $p_i$ . After ‘eliminating’  $P_i, P_{i+1}$  we obtain again a non-crossing partition (we relabel the indices accordingly). We can repeat this elimination procedure, say for  $r$  steps, until we arrive at (again after relabeling)  $(Q_1, Q_2, \dots, Q_{k-2r})$ , with  $Q_i = (q_i, q_{i+1})$  and a pair partition  $\pi'$  on  $\{1, \dots, k-2r\}$ , where all values of  $n_i$  associated to partition blocks  $\{i, i+1\}$  strictly larger than zero. We can restore  $P_1, \dots, P_k$  from  $(Q_1, Q_2, \dots, Q_{k-r})$  by  $N^r$  choices. Let  $\{i, i+1\}$  be a partition block of  $\pi'$  associated to some  $n_j > 0$  (such a block exists as  $n_1 > 0$ ). As before we choose  $q_i, q_{i-1}, q_{i-2}, \dots, q_{i+2}$ , for which there are  $CN^{1+\frac{k}{2}-r-1}$  possibilities ( $N$  possibilities for  $q_i$ , then  $N$  possibilities for each of the  $\frac{k}{2} - r$  equivalence classes except the equivalence class  $\{i, i+1\}$ ). For  $q_{i+1}$  there are by Assumption 2.2 only  $o(N)$  possibilities. Together with the  $N^r$  possibilities to obtain  $P_1, \dots, P_k$  from  $(Q_1, Q_2, \dots, Q_{k-r})$ , this proves the claim in (4.36) and hence completes the proof of statement (i).

We recall, that statement (ii) reads

$$\lim_{N \rightarrow \infty} \frac{1}{N^{k+2}} \sum_{P \in S_k, Q \in S_k} \mathbb{E}[X_N(P)X_N(Q)] - \mathbb{E}[X_N(P)]\mathbb{E}[X_N(Q)] = 0.$$

The main idea is to use (2.8), which is a consequence of (2.3) and (2.4). Hence we introduce the following notation for  $P, Q \in S_k$

$$d(P, Q) := \min_{1 \leq i, j \leq k} \|P_i - Q_j\|_{\varphi_N}.$$

Further we write  $[P, Q] := (P_1, \dots, P_k, Q_1, \dots, Q_k)$  and for a pair partition  $\pi \in \mathcal{PP}(2k)$ , we say that  $[P, Q] \in S_k \times S_k$  with  $(l_1, \dots, l_{2k}) = G_N([P, Q])$  (see (4.30)) is  $\pi$ -consistent (we write  $[P, Q] \in S_{2k}(\pi)$ ), if

$$i \sim_{\pi} j \Leftrightarrow \{\varphi_N^{(-1)}[P, Q]_i, \varphi_N^{(-1)}([P, Q]_j)\} = \{l_n, l_{n+1}\} \quad \text{for some odd } n.$$

Analogously to the definition of  $m$  in the proof of (i), for  $n_1, \dots, n_{2k}, \delta \in \mathbb{N}$  and  $\pi \in \mathcal{PP}(2k)$ , we set

$$\begin{aligned} M(n_1, n_3, \dots, n_{2k-1}, \delta, \pi) := \\ \#\{[P, Q] \in S_{2k}(\pi) : (G_N([P, Q]))_{j+1} - (G_N([P, Q]))_j = n_j, j = 1, 3, 5, \dots \\ \text{and } d(P, Q) = \delta\} \end{aligned}$$

By (2.8) there are constants  $C > 0$  and  $\beta \in [0, 1)$  such that

$$\begin{aligned} & \left| \frac{1}{N^{k+2}} \sum_{P \in S_k, Q \in S_k} \mathbb{E}[X_N(P)X_N(Q)] - \mathbb{E}[X_N(P)]\mathbb{E}[X_N(Q)] \right| \\ & \leq C \frac{1}{N^{k+2}} \sum_{\pi \in \mathcal{PP}(2k)} \sum_{n_1, n_3, \dots, \delta \geq 0} M(n_1, n_3, \dots, n_{2k-1}, \delta, \pi) \beta^{n_1 + n_3 + \dots + n_{2k-1} + \delta}. \end{aligned}$$

Again, by the finiteness of the geometric series, it suffices to show that for any  $n_1, n_3, \dots, n_{2k-1}, \delta, \pi$

$$M(n_1, n_3, \dots, n_{2k-1}, \delta, \pi) = o(N^{k+2}). \quad (4.37)$$

For given  $n_1, n_3, \dots, n_{2k-1}, \delta, \pi$ , we estimate the number of elements  $[P, Q]$  that contribute to  $M(n_1, n_3, \dots, n_{2k-1}, \delta, \pi)$  for which  $\delta = d(P, Q) = \|P_{i'} - Q_{j'}\|_{\varphi_N}$ , that means the minimal distance is obtained between  $P_{i'}$  and  $Q_{j'}$  (the number of such possible indices  $i'$  and  $j'$  depends on  $k$  only and may hence be neglected for the consideration of  $N \rightarrow \infty$ ). When choosing  $[P, Q]$  we proceed in the following order: We start with  $P_{i'}$  and first fix all values  $P_{i'+1}, \dots, P_k, P_1, \dots, P_{i'-1}$ . Then we proceed with  $Q_{j'}$  and the remaining values of  $Q$ , i.e.  $Q_{j'+1}, \dots, Q_k, Q_1, \dots, Q_{j'-1}$ . As before, we have  $N^2$  choices for the starting point  $P_i$  and  $N$  choices whenever we encounter a 'new' equivalence class, otherwise there is only a constant number of choices. There are  $k$  equivalence classes, resp.  $k-1$  'new' equivalence classes, once  $P_{i'}$  is fixed. Here it is crucial, that when we fixed all values of  $P$ , we have only a constant number of choices for  $Q_{j'}$  (by the restriction  $\delta = \|P_{i'} - Q_{j'}\|_{\varphi_N}$ ), i.e. the  $CN^2$  possibilities for  $(P_{i'}, Q_{j'})$  provide a starting point for both  $P$  and  $Q$  rather than having  $N$  further possibilities for  $Q_{j'}$ . This shows (4.37) and hence completes the proof of the main theorem.  $\square$

It remains to prove Proposition 2.8, which states that under certain conditions on the mixed moments of the stochastic process, the fourth moment of the trace does not converge to the fourth moment of the semicircle law, i.e. we show that for some  $C > 0$  and  $N$  large enough we have

$$\mathbb{E} \left( \frac{1}{N} \text{tr} A_N^4 \right) \geq \kappa_2 + C.$$

*Proof of Proposition 2.8.* We need to estimate

$$\frac{1}{N^3} \sum_{P \in S_4} \mathbb{E} [X_N(P_1)X_N(P_2)X_N(P_3)X_N(P_4)]$$

from below. As in the proof of Lemma 4.1, terms of the form  $\mathbb{E}(X_N(P_1)^4)$  give a vanishing contribution to the sum as  $N \rightarrow \infty$ . Recall that by the assumption of the proposition we have

$$\mathbb{E}(Z_{i_1}Z_{i_2}Z_{i_3}Z_{i_4}) = C\beta^{n_1+n_3}, \quad \mathbb{E}(Z_{i_1}^2Z_{i_2}^2) = 1$$

for some  $0 < \beta < 1$ . Hence, as again  $\frac{\#\{P: P \in S_4^0\}}{N^3} \rightarrow \kappa_2$  for  $N \rightarrow \infty$ , we obtain for any  $\varepsilon > 0$  and  $N$  sufficiently large

$$\begin{aligned} & \frac{1}{N^3} \sum_{P \in S_4} \mathbb{E} [X_N(P_1)X_N(P_2)X_N(P_3)X_N(P_4)] - \kappa_2 \\ & \geq C \frac{1}{N^3} \sum_{\pi \in \mathcal{PP}(k)} \sum_{n_1, n_3 \geq 0, \exists i: n_i \neq 0} m(n_1, n_3, \pi) \beta^{n_1+n_3} - \varepsilon \\ & \geq C \frac{1}{N^3} m(1, 1, \pi') \beta^2 - \varepsilon \end{aligned} \tag{4.38}$$

for  $\pi' = \{\{1, 2\}, \{3, 4\}\}$ . Observe that the estimate in (4.38) is justified as  $\beta > 0$ . We have to estimate (from below) the number of possible choices for  $p_1, \dots, p_4$  with  $\|P_1 - P_2\|_{\varphi_N} = \|P_3 - P_4\|_{\varphi_N} = 1$  (see Fig. 3). We will show

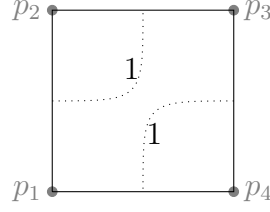


FIGURE 3. The number of possible values for  $p_1, \dots, p_4 \in \{1, \dots, N\}$  such that  $\|P_1 - P_2\|_{\varphi_N} = \|P_3 - P_4\|_{\varphi_N} = 1$  (indicated by the dotted lines) is larger than  $CN^3$  for some  $C > 0$ .

that this number is larger than  $CN^3$  for some positive constant  $C$ . We can assume that  $p_3 - p_1 = 1$  (this further reduces the number of choices). Then  $P_1, P_2$  resp.  $P_3, P_4$  are neighbors according to Definition 2.1. We set

$$\begin{aligned} x_i &:= \#\{j : (\varphi_N(j))_2 = i, (\varphi_N(j+1))_2 = i+1 \text{ or} \\ &\quad (\varphi_N(j))_2 = i+1, (\varphi_N(j+1))_2 = i\} \\ y_i &:= \#\{j : (\varphi_N(j))_1 = i, (\varphi_N(j+1))_1 = i+1 \text{ or} \\ &\quad (\varphi_N(j))_1 = i+1, (\varphi_N(j+1))_1 = i\}. \end{aligned}$$

Here,  $x_i$  is the number of horizontal neighbors in the  $i$ -th and  $i+1$ -st column (in the upper triangular matrix) with distance one (according to  $\|\cdot\|_{\varphi_N}$ ). Analogously,  $y_i$  is the number of vertical neighbors in the  $i$ -th and  $i+1$ -st row (in the upper triangular matrix) with distance one (according to  $\|\cdot\|_{\varphi_N}$ ). For each choice of  $p_1, p_3$ , say  $p_1 = i, p_3 = i+1$ , we have  $(x_i + y_i)^2$  possibilities for  $(p_2, p_4)$ . We observe that by the assumption of the lemma we have

$$\sum_{i=1}^{N-1} x_i + y_i \geq CN^2.$$

By the Cauchy-Schwarz inequality we have

$$m(1, 1, \pi') \geq \sum_i (x_i + y_i)^2 \geq \frac{1}{N} \left( \sum_i x_i + y_i \right)^2 \geq CN^3.$$

Inserting this estimate into (4.38) proves the claim.  $\square$

## REFERENCES

- [1] Anderson, G.W., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices. Cambridge studies in advanced mathematics 118. Cambridge University Press, Cambridge (2010)
- [2] Arnold, L.: On Wigner's semicircle law for the eigenvalues of random matrices. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **19**, 191–198 (1971)
- [3] Bryc, W., Dembo, A., Jiang, T.: Spectral measure of large random Hankel, Markov and Toeplitz matrices. Ann. Probab. **34**(1), 1–38 (2006). DOI 10.1214/009117905000000495. URL <http://dx.doi.org/10.1214/009117905000000495>

- [4] Doob, J.L.: The Brownian movement and stochastic equations. *Ann. of Math. (2)* **43**, 351–369 (1942)
- [5] Erdős, L.: Universality of Wigner random matrices: a survey of recent results. *Uspekhi Mat. Nauk* **66**(3(399)), 67–198 (2011). DOI 10.1070/RM2011v066n03ABEH004749. URL <http://dx.doi.org/10.1070/RM2011v066n03ABEH004749>
- [6] Erdős, L., Schlein, B., Yau, H.T.: Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.* **287**(2), 641–655 (2009). DOI 10.1007/s00220-008-0636-9. URL <http://dx.doi.org/10.1007/s00220-008-0636-9>
- [7] Friesen, O., Löwe, M.: On the limiting spectral density of symmetric random matrices with correlated entries. In: *Random matrices and iterated random functions, Springer Proc. Math. Stat.*, vol. 53, pp. 3–29. Springer, Heidelberg (2013). DOI 10.1007/978-3-642-38806-4\_1. URL [http://dx.doi.org/10.1007/978-3-642-38806-4\\_1](http://dx.doi.org/10.1007/978-3-642-38806-4_1)
- [8] Friesen, O., Löwe, M.: A phase transition for the limiting spectral density of random matrices. *Electron. J. Probab.* **18**, no. 17, 17 (2013). DOI 10.1214/EJP.v18-2118. URL <http://dx.doi.org/10.1214/EJP.v18-2118>
- [9] Friesen, O., Löwe, M.: The semicircle law for matrices with independent diagonals. *J. Theoret. Probab.* **26**(4), 1084–1096 (2013). DOI 10.1007/s10959-011-0383-2. URL <http://dx.doi.org/10.1007/s10959-011-0383-2>
- [10] Georgii, H.O.: *Gibbs Measures and Phase Transitions*. De Gruyter Studies in Mathematics 9. Walter de Gruyter, Berlin (1988)
- [11] Götze, F., Tikhomirov, A.N.: Limit theorems for spectra of random matrices with martingale structure. *Teor. Veroyatn. Primen.* **51**(1), 171–192 (2006). DOI 10.1137/S0040585X97982268. URL <http://dx.doi.org/10.1137/S0040585X97982268>
- [12] Guionnet, A.: *Large random matrices: lectures on macroscopic asymptotics, Lecture Notes in Mathematics*, vol. 1957. Springer-Verlag, Berlin (2009). DOI 10.1007/978-3-540-69897-5. URL <http://dx.doi.org/10.1007/978-3-540-69897-5>. Lectures from the 36th Probability Summer School held in Saint-Flour, 2006
- [13] Hochstättler, W., Kirsch, W., Warzel, S.: Semicircle law for a matrix ensemble with dependent entries. preprint, to appear in *Journal of Theoretical Probability* (2014)
- [14] Isserlis, L.: On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika* **12**(1/2), 134–139 (1918). URL <http://www.jstor.org/stable/2331932>
- [15] Levin, D.A., Peres, Y., Wilmer, E.L.: *Markov chains and mixing times*. American Mathematical Society (2006)
- [16] Schenker, J., Schulz-Baldes, H.: Semicircle law and freeness for random matrices with symmetries or correlations. *Math. Res. Lett.* **12**, 531–542 (2005)
- [17] Schubert, K.: Spectral density for random matrices with independent skew-diagonals. arXiv:1510.06448, Preprint, submitted (2015)
- [18] Tao, T., Vu, V.: Random matrices: universality of ESDs and the circular law. *Ann. Probab.* **38**(5), 2023–2065 (2010). DOI 10.1214/10-AOP534. URL <http://dx.doi.org/10.1214/10-AOP534>. With an appendix by Manjunath Krishnapur
- [19] Tao, T., Vu, V.: Random matrices: Universality of local eigenvalue statistics. *Acta Mathematica* **206**, 127–204 (2011). URL <http://dx.doi.org/10.1007/s11511-011-0061-3>
- [20] Wigner, E.P.: Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math. (2)* **62**, 548–564 (1955)
- [21] Wigner, E.P.: On the distribution of the roots of certain symmetric matrices. *Ann. of Math.* **67**, 325–328 (1958)

INST. FOR MATH. STAT., UNIV. MÜNSTER, ORLÉANS RING 10, 48149 MÜNSTER,  
GERMANY

*E-mail address:* `maloewe@uni-muenster.de`

INST. FOR MATH. STAT., UNIV. MÜNSTER, ORLÉANS RING 10, 48149 MÜNSTER,  
GERMANY

*E-mail address:* `kristina.schubert@uni-muenster.de`